Dense Heterogeneous Integration for InP Bi-CMOS Technology

Y. Royter, P.R. Patterson, J.C. Li, K.R. Elliott, T. Hussain, M.F. Boag-O'Brien, J.R. Duvall, M.C. Montes, D.A. Hitko, J.S. Sewell*, M. Sokolich, D.H. Chow, P.D. Brewer
HRL Laboratories LLC, Malibu, CA, 90265, USA; Phone: (310) 317-5548; Email: yroyter@hrl.com
*Air Force Research Laboratory, Wright-Patterson AFB, OH, 45433, USA

ABSTRACT

InP Bi-CMOS technology capable of wafer-scale device-level heterogeneous integration (HI) of InP HBTs and CMOS has been developed. With this technology, full simultaneous utilization of III-V device speed and CMOS circuit complexity is possible. Simple ICs and test structures have been fabricated, showing no significant CMOS or HBT degradation and high heterogeneous interconnect yield. The heterogeneously integrated differential amplifiers with record performance and HBTs with f_T=400GHz were obtained. Thermal vias to the Si substrate provide sufficient heat path to lower HI HBT thermal resistances close to on-InP values. Resulting circuits maintain maximum CMOS integration density and HBT performance, while keeping the heterogeneous interconnect length below 5μm.

INTRODUCTION

Despite much higher level of investment in CMOS, III-V compound semiconductor (CS) device performance still surpasses that of their Si counterparts. However, this investment has produced CMOS ICs with orders of magnitude higher integration level at a fraction of the cost. For many mixed signal applications, having circuits composed of both Si CMOS, which possesses low power dissipation and high transistor count, and compound semiconductor transistors with high-speed high-voltage swing performance would be advantageous. In general, heterogeneous integration (HI) of technologies from dissimilar materials can be done in a variety of ways. The heterogeneous integration method depends primarily on the final circuit or system. Such considerations as whether the different technology components are integrated as full circuits, circuit blocks or as individual devices, as well as the required heterogeneous interconnect impedance, will determine the required interconnect density and length. Systems in a package or systems on a chip consisting of separate circuits can be interconnected with wire bonding if the number of interconnections is small and the interconnect signal speed is not extremely high. For systems on a chip where interconnect number is large, such as a focal plane array integrated with a readout, or where wire bond impedance is prohibitive, bump bonding is more advantageous. For circuits that require device-level integration where a large portion of devices from one material are connected to a device from another material, heterogeneous interconnects comparable in length and density to standard semiconductor technology interconnects are required.

One way to achieve device-level integration is by performing metamorphic epitaxial growth, such as InGaAs on Si [1], and then processing both types of devices together. Another approach is to process the first device type before the growth and processing of the second device type. This can also be done after the full interconnect stack for the first device type is finished by etching windows in the dielectric before growth [2]. Instead of epitaxial growth, wafer bonding can be utilized for heterogeneous integration of two materials. The bonding can be done on unprocessed wafers, such as done for some SOI wafer fabrication, or with fully or partially processed wafers. The choice of the process is determined by such considerations as process compatibility between the two device types, required heterogeneous interconnect length, heat removal, reliability, and cost. A well designed process, thus, would produce ICs where device performance, reliability, size, thermal management and minimum interconnect length of both device types would be on par with non-HI technologies.

Device-level integration of compound semiconductor devices with CMOS would enable a new class of high-performance high-functionality ICs [3, 4] at a cost similar to CS ICs alone. For such heterogeneous ICs, InP DHBT technology is particularly suited due to the high device performance [5], high IC speed at low power [6], and circuit compactness as illustrated in Figure 1. The heterogeneous integration approach can use the latest CMOS technology, where as a monolithically integrated technology, such as SiGe Bi-CMOS, lags in the CMOS technology node and incorporation of the state-of-the-art SiGe HBTs is impeded by small market size relative to the development cost. In this

Technology	Speed	Power	TX Count	Market Driver	Pros/ Cons
InP DHBT	High	Mod	$<10^4$	Defense	BJT only, functionality limited
CMOS	Mod	Low	$>10^9$	Commercial	No precision fast device, low drive
SiGe HBT	Mod	Mod	$<10^6$	Commercial Defense	Lags latest CMOS
CoSMOS	High	Low	$<10^4$ / $>10^9$	Commercial Defense	Fastest CMOS and HBTs

Figure 1. Comparison of transistor speed, transistor count, power dissipation and target market of various technologies. HRL's CoSMOS heterogeneous integration technology can combine advantages of all these technologies for many circuit applications.

978-1-4244-3433-6/09/$25.00©2009 IEEE

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAY 2009		**00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Dense Heterogeneous Integration for InP Bi-CMOS Technology	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory,Wright-Patterson AFB,OH,45433**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002192. Presented at the IEEE International Conference on Indium Phosphide and Related Materials (IPRM 2009) Held in Newport Beach, California on May 10-14, 2009. Sponsored by ONR.

14. ABSTRACT
InP Bi-CMOS technology capable of wafer-scale device-level heterogeneous integration (HI) of InP HBTs and CMOS has been developed. With this technology, full simultaneous utilization of III-V device speed and CMOS circuit complexity is possible. Simple ICs and test structures have been fabricated, showing no significant CMOS or HBT degradation and high heterogeneous interconnect yield. The heterogeneously integrated differential amplifiers with record performance and HBTs with fT=400GHz were obtained. Thermal vias to the Si substrate provide sufficient heat path to lower HI HBT thermal resistances close to on-InP values. Resulting circuits maintain maximum CMOS integration density and HBT performance, while keeping the heterogeneous interconnect length below 5μm.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**6**	

work, we describe InP Bi-CMOS technology with wafer-scale device-level heterogeneous integration of 250nm, 300GHz f_T/f_{MAX} InP DHBTs with IBM's 130nm RF-CMOS (CMRF8SF).

PROCESS TECHNOLOGY

For this technology, InP DHBT epitaxial layers are bonded to the top interconnect layer of a CMOS wafer, which are then processed into HBTs and interconnected with the CMOS [7]. This process is illustrated in Figure 2 and

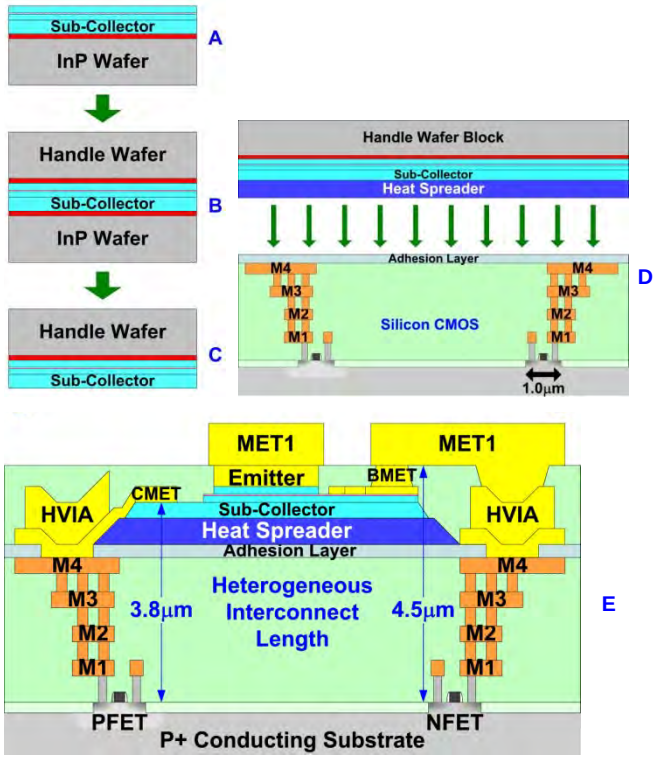

Figure 2. Schematic representation of heterogeneous integration process consisting of a) HBT epitaxial growth, b) temporary wafer bonding to a handle wafer, c) growth substrate removal, d) permanent bonding to CMOS after heat spreader deposition, and e) HBT fabrication and heterogeneous integration.

images of the wafer at different stages are shown on Figure 3. The IBM CMOS process is terminated at the last planar interconnect layer with only a thin dielectric layer on top of the copper. All of the CMOS interconnect layers, except for the last one, are planarized using chemical mechanical polishing (CMP), so the process can be stopped at any of these layers. The CMP surface is planar enough for wafer bonding and enables high yield epitaxial transfer. The 200mm CMOS wafers are then cut into four 3inch wafers so that they can be further processed in our HBT line. Standard DHBT epitaxial layers are grown with the emitter side up on an InP substrate with thin etch stop layers inserted between the subcollector and the InP substrate (Figure 2a). The DHBT wafer is then wafer bonded to a temporary Si handle wafer (Figure 2b). The InP growth substrate is removed by wet etching, stopping on the etch stop layers (Figure 2c). An Al heat spreader layer is then deposited on the subcollector. The epitaxial layers are then permanently bonded to the top CMOS surface using BCB adhesive layer (Figure 2d) [8]. The permanent bond can be a full-wafer wafer bond. However, die bonding was developed to allow the use of different size InP growth substrate and the substrate on which the HBT process is carried out. Thus, the HBT process can be carried out on the full size CMOS wafer. After the permanent bonding, the handle is removed, completing the epitaxial transfer process. The epitaxial transfer process consists of two bonding steps to ensure optimal epitaxial growth and device layout, which is with the emitter up in both cases.

After the epitaxial transfer, the InP HBT process is carried out. The CMOS wafer contains the alignment marks needed for the HBT processing. InP DHBT fabrication process on top of CMOS is similar to the standard process [9]. Alignment accuracy to the CMOS and between the HBT layers is limited by the stepper and is equal to our standard HBT process. Despite increased overall topography, critical dimension control for both electron beam and optical lithography is maintained. This alignment strategy enables dense and intimate device-level integration with <5μm heterogeneous interconnect length. After the HBT is fabricated, the heat spreader and the adhesion layers are patterned. The HBT and CMOS are finally interconnected through HI vias which connect the HBT interconnect metal to the top CMOS Cu interconnect layer (Figure 2e). In addition to electrical vias, thermal vias that connect the heat spreader of each HBT to the Si substrate are also fabricated. Only one post-CMOS interconnection layer is used. All other

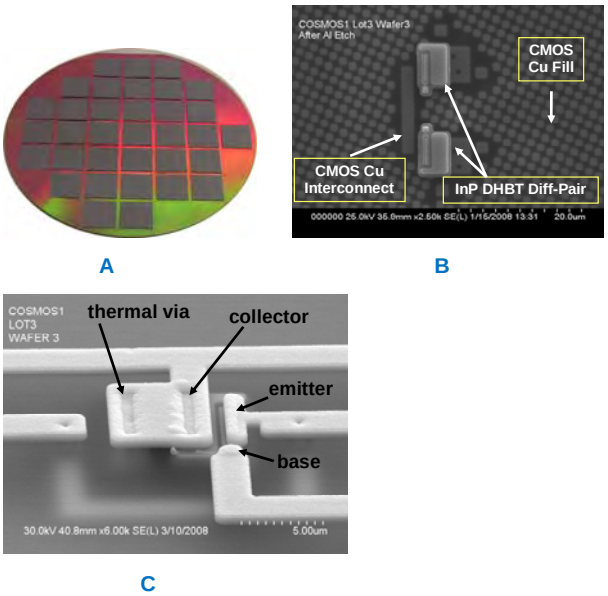

Figure 3. a) A photograph of an die bonded wafer at step E of Figure 2, b) an SEM micrograph HI integrated HBT before final interconnection, and c) an SEM micrograph of an HBT at the end of the process.

interconnect layers as well as resistors and capacitors are fabricated as part the CMOS interconnect process, so the current cost of the ICs is similar to our standard DHBT process.

RESULTS

To verify that the IBM CMOS devices are unaffected by the heterogeneous integration process, performance of the CMOS test devices that have gone through the IBM process only and ones that also went through the heterogeneous integration were compared. The IV characteristics of various IBM CMRF8SF library devices from HRL heterogeneously integrated wafers showed no signs of device degradation when compared to control IBM CMOS wafers. Figure 4

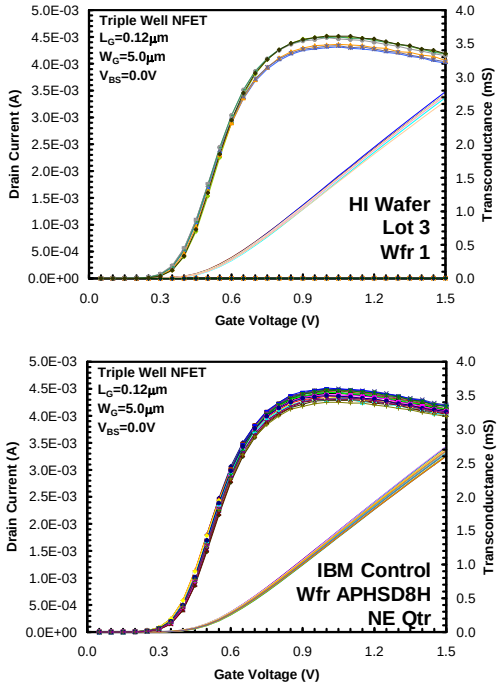

Figure 4. Transfer IV characteristics and associated transconductance, gm, from a 1.5V thin-oxide, triple-well NFET (LG=120nm, WG=5.0µm) bias at a VDS=1.5V and VBS=0.0V. A sparse sampling of 10 devices from a single HI wafer (top) and 22 devices from a quarter of the IBM control CMOS wafer (bottom) are shown.

shows the typical transfer characteristics of a 1.5V thin-oxide, triple well NFET sampled from an HRL fabricated HI wafer and an IBM fabricated control CMOS wafer.

The HI HBTs performance was on par with our standard 250nm HBTs. For an A_E=0.25x4µm^2 InP DHBT, Figure 5 (top) shows the forward Gummel characteristics and a DC current gain (β_f) >30; Figure 5 (bottom) shows common emitter IV characteristics with low output conductance. For the same size device, Figure 6 shows f_T and f_{MAX} values of 400GHz and 244GHz, respectively. Variation in f_T is also shown, with > 90% of the device f_T values > 300GHz.

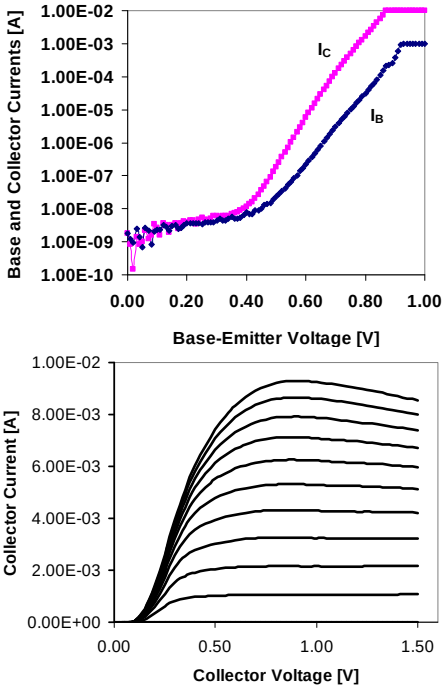

Figure 5. DC characteristics of an A_E=0.25x4.0µm^2 InP DHBT bonded onto CMOS: (top) forward Gummel with DC current gain (β_f) >30 while being driven at an I_C=10mA/µm^2 and a V_{CE}=V_{BE}; (bottom) common emitter IV and corresponding V_{BE} with I_B=0 to 200µA in 20µA steps.

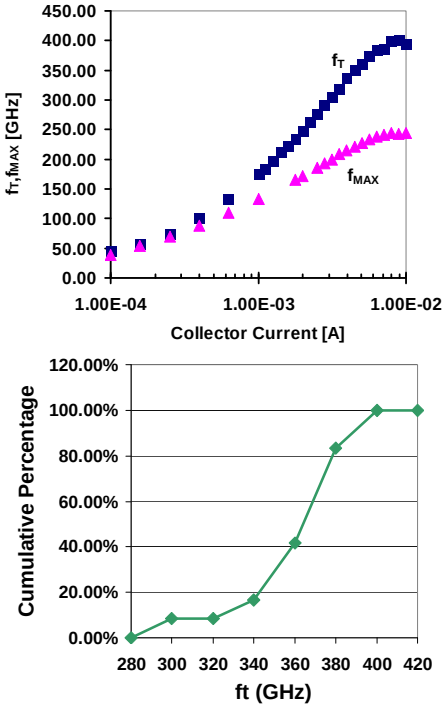

Figure 6. The f_T and f_{MAX} dependence with I_C of an A_E=0.25x4.0µm^2 InP DHBT while being driven by a forced I_B and fixed V_{CE}=1.2V (top). Cumulative distribution of the for devices across a wafer (bottom) shows that high performance is not limited to select devices.

Since the HBTs are located on the interconnect stack dielectric, a heat path is provided through a thermal via. A cross sectional SEM of the HBT and its thermal vias is shown on Figure 7. This heat path connects the HBT

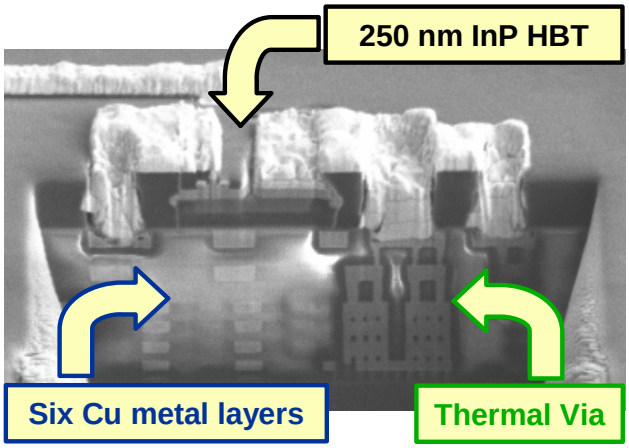

Figure 7. Cross sectional SEM showing the HBT with its collector connected through a thermal via to the Si substrate.

collector contact to the Si substrate through the HI metal, large HI via and the CMOS copper. An aluminum heat spreader layer underneath the sub-collector lowers the thermal resistance from the base-collector junction to the collector contact. Two types of thermal vias were made: DC via with continuous metal contact and AC via with a thin dielectric layer between the HI metal and the copper interconnect. AC via reduces the parasitic capacitance associated with the thermal via [10]. Thermal resistance (R_{TH}) of the heterogeneously integrated HBTs was determined by measuring the V_{BE} shift as a function of dissipated power at various ambient temperatures [11]. Thermal resistances for devices with DC via, AC via, and no thermal via are shown on Figure 8. Thermal resistance values as low as 12.0 °C/mW were obtained, which are close to 8-10 °C/mW obtained for standard HBTs [12]. As can be seen when comparing to devices without the via, the thermal via greatly reduces HBT thermal resistance. Furthermore, R_{TH} values for DC and AC thermal vias are not significantly different, so the AC via devices with lower parasitic capacitance can be used in the circuits. It should be noted that the device in Figure 6 has a DC thermal via, and devices with AC thermal via show very similar performance.

In order to assess the yield and resistive loss of the HI via, a 1000 unit long chain of the nominal 1.0x1.0μm^2 HI vias at a pitch of 5μm were electrically tested. These vias connect the top CMOS copper metal to the HBT interconnect metal. Figure 9 shows the average HI via resistance is <200mΩ across the majority of the 3inch diameter wafer. Both stand alone via chain test structures (top) and via chains with HBTs 1μm away from the via (bottom) show very similar via resistance, demonstrating possibility for high integration density in this technology.

Technology robustness was assessed by comparing differential amplifier performance before and after wafer temperature cycling in an inert gas ambient. The temperature was cycled from -55°C to +85°C with 10 min dwell time at extremes. Comparison of slew rate and DC Gain x Unity-Gain-Bandwidth figures of merit before and after 100 thermal cycles (not under bias) is shown on Figure 10. This data indicates that our process produces robust devices and circuits.

Using the InP Bi-CMOS technology, differential amplifier ICs [10] and DC and RF test structures were fabricated. The differential amplifier shown on Figure 11 demonstrated gain bandwidth product of 40-130 GHz, low frequency gain of >48dB, and a DC Gain x Unity-Gain-Bandwidth figure of merit of 1.6x10^4 GHz . The use of InP DHBTs supports a 6.9 V differential output swing with 0.4V differential input. A slew rate of 4.4x10^4 V/μs is achieved with as low as 40mW dissipated power. Presence of CMOS enabled the use of novel on-chip buffer circuits to facilitate the on-wafer characterization of these amplifiers. As can be seen from the amplifier schematic, intimate device-level heterogeneous integration had to be achieved for its realization.

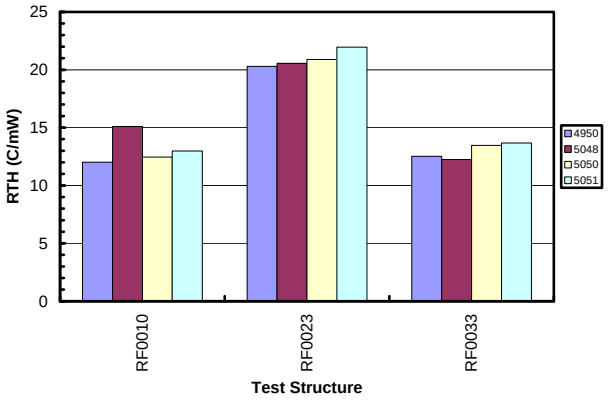

Figure 8. Measured HBT thermal resistance across a wafer (4 reticles) for three device types: RF0010 with DC thermal via, RF0023 with no thermal via, and RF0023 with AC thermal via.

CONCLUSIONS

Novel InP Bi-CMOS technology with wafer-scale device-level heterogeneous integration is demonstrated. Dense HI integration without device degradation enables robust ICs that maintain both CS performance and CMOS functionality. Differential amplifier ICs in this technology show record performance not possible with either technology alone. This integration approach is applicable to any generation CMOS and HBTs, as well as other CS devices. CS device integration with CMOS could accelerate its maturity and utilization.

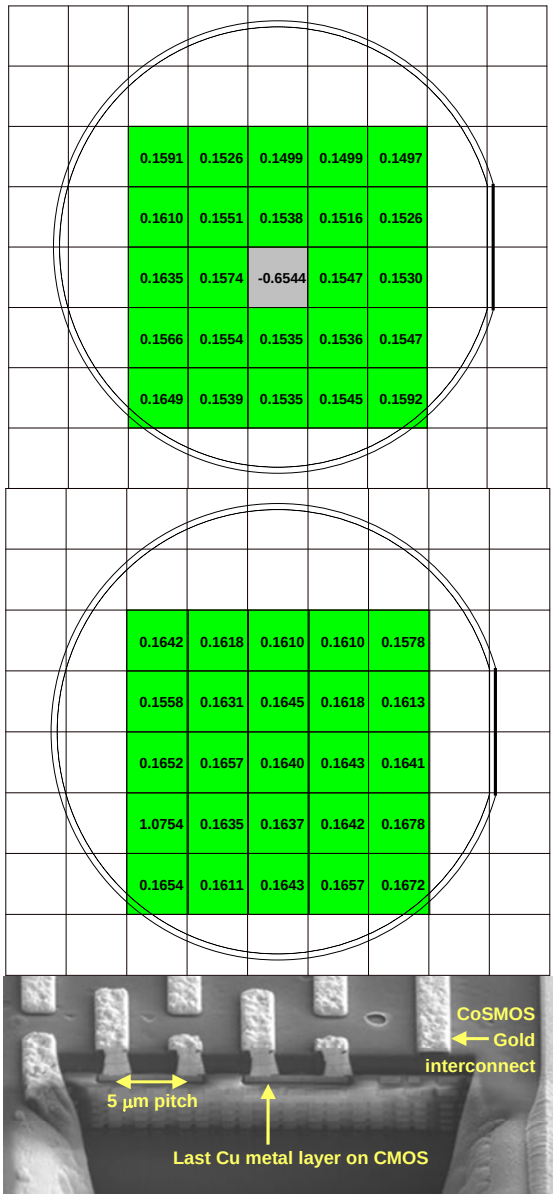

Figure 9. Wafer map showing the unit resistance in ohms of a 1000 unit long heterogeneous interconnect via chain with a 5µm pitch and 1.0x1.0µm^2 via size. Top map is for stand alone vias and bottom map is for vias with HBTs placed in 1.0µm proximity. Fully processed wafer and a cross sectional SEM of the via chain are also shown.

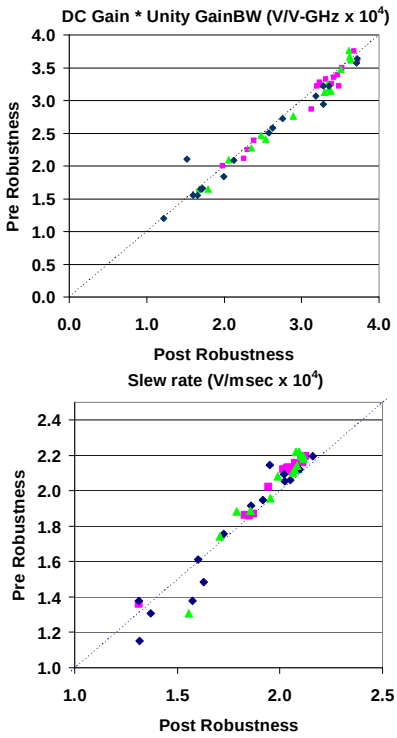

Figure 10. HI differential amplifier slew rate (bottom) and DC gain x unity gain bandwidth (top) comparison before and after 100 of -55°C to +85°C temperature cycles, indicating that the HI ICs are robust.

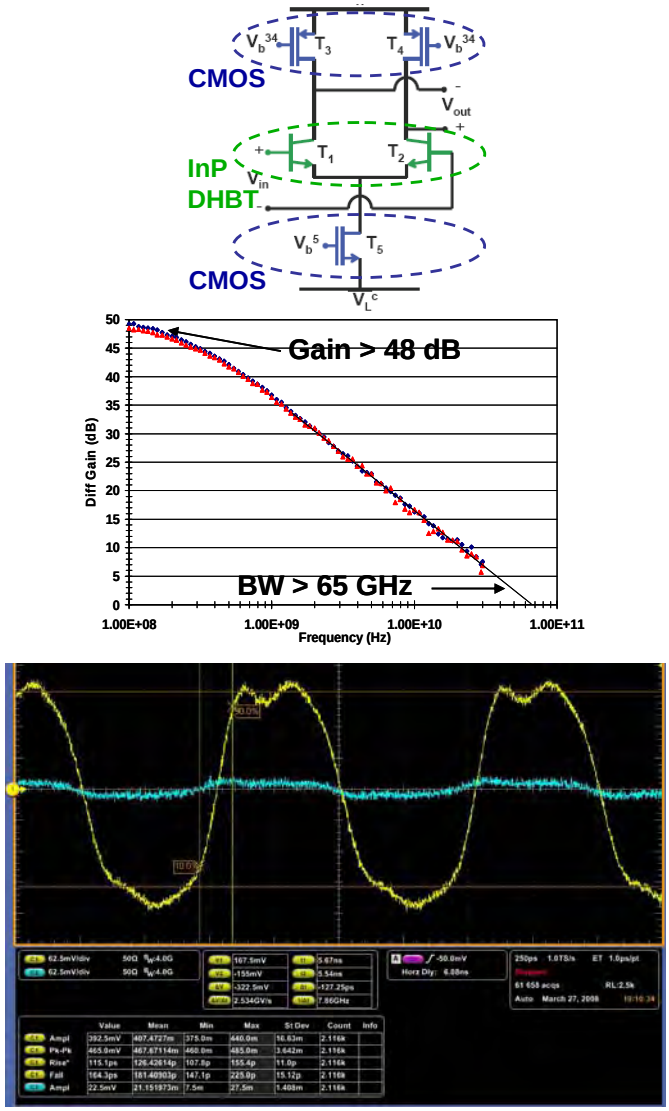

Figure 11. Heterogeneously integrated differential amplifier circuit schematic (top), measured gain versus frequency (middle) and digital oscilloscope capture of the measured large signal output swing (bottom). Low frequency gains of 48 dB were achieved simultaneously with a unity gain bandwidth >65 GHz. A 10-90% slew rate of 4.4×10^4 V/µs is observed with a 6.9V differential output voltage swing and a 0.4V differential input voltage swing.

ACKNOWLEDGEMENTS

This work was supported in part by DARPA (Dr. Mark Rosker, PM) through AFRL (Dr. Gregory Creech) contract FA8650-07-C-7714 (CoSMOS).

REFERENCES

[1] M. Hudait et al., 2007 IEDM Tech. Dig., p. 625
[2] J.F. Ahadian et al., IEEE J. Quantum Electron., v. 34(7), p. 1117 (1998)
[3] M.J. Rosker et al, 2008 CSICS Proceedings, p. 6
[4] K.R.Elliott et al, 2008 MRS Spring Meeting, p. 1068-C01-06
[5] M. Feng et al, 2007 IPRM Proceedings, p. 399
[6] T. Hussain et al, 2006 IPRM Proceedings, p. 85
[7] J.C. Li et al, 2008 IEDM Proceedings, p. 944
[8] P.R. Patterson et al, 2008 ECS Trans., v.16., i. 8, p. 221
[9] T. Hussain et al, 2004 IEDM Technical Digest, p.553
[10] J.C. Li et al, 2008 CSICS Proceedings, p. 227
[11] D.E. Dawson et al, IEEE Trans. Elec. Dev., v. 39(10), p.2235
[12] J.C. Li et al, 2005 CSICS Proceedings, p. 65